Scaling Multiple Point Statistics for Non-Stationary Geostatistical Modeling

Julián M. Ortiz, Steven Lyster and Clayton V. Deutsch

Centre for Computational Geostatistics
Department of Civil & Environmental Engineering
University of Alberta

Multiple-point statistics are used in geostatistical simulation to improve forecasting of responses that are highly dependent on the reproduction of complex features of the phenomenon that cannot be captured by conventional two-point simulation methods. Inference of multiple-point statistics is often based on a training image that depicts the features that provide the character to the geological event being modelled. One limitation of this approach is that the univariate distribution of categories (facies or rock types) in the training image may not match the target statistics of the final model. The question of scaling multiple-point statistics arises, the idea being to take the statistics from the training image and scale them in a repeatable manner and honouring the target univariate proportions of categories.

An iterative scaling approach based on the expression for scaling multiple-point statistics in a purely random case is proposed. A multi-Gaussian alternative is discussed, which shows some problems due to the symmetry of the multivariate Gaussian distribution. The implementation is illustrated through an example where it is shown that the proposed method lies in between two extreme cases for a Boolean simulation, namely, the change in size of the objects and the change in their number of occurrences. A second example is presented to illustrate the potential use of this scaling procedure for non-stationary multiple-point geostatistical simulation.

Introduction

Multiple-point statistics can be used in an attempt to improve geostatistical modelling of variables distributed in space. The relationships between several points at a time are estimated from training data and imposed during the simulation process, in order to achieve a numerical model that correctly represents these intricate relationships that conventional simulation cannot capture.

Figure 1 shows a reference image where the relationships between the different geological units or facies cannot be correctly captured by conventional simulation methods. Exhaustive images such as the one depicted here can be used as an analogue to the phenomenon that is being modelled; multiple-point statistics can be extracted by scanning the image and computing the frequency with which facies arranged in specific patterns occur.

The training information requires abundant data located over a regular grid of points, in order to have enough replicates of each particular multiple-point event. This is often solved by utilizing a training image (Guardiano and Srivastava, 1992; Deutsch, 1992) or by using available pseudo-regularly spaced production data, as is the case in mining applications (Ortiz, 2003; Ortiz and Deutsch, 2004; Ortiz and Emery, 2005).

Multiple-point geostatistical simulation can be performed using any of the available methods. The single normal equation simulation proposed by Strebelle and Journel (2000) estimates the

conditional distribution at every location given a multiple-point configuration by calculating the frequency with which the indicator at the location being simulated is one given that the multiple-point event occurs in the training image. Alternatively, simulated annealing (Deutsch, 1992) can be used to match the multiple-point frequencies extracted from a training image into a simulated numerical model. Again, the multiple-point statistics are read from the training image as frequencies of particular events occurring. Other methods such as neural networks also rely on the use of a training image to extract and reproduce the multiple-point statistics (Caers and Journel, 1998).

There is an implicit assumption of stationarity during the exporting of multiple-point statistics from the training image to the simulated model. A stationarity assumption stricter than the usual second order stationarity is required, since the use of configurations of several points also locks all lower order statistics. For instance, if four-point configurations are matched during the multiple-point simulation process, then all three-point and two-point statistics whose configuration is included in the four-point configuration originally used are implicitly matched (

Figure 2). Most importantly, the histogram (one-point statistic) is also locked when higher order statistics are honoured. This paper addresses the issue of scaling multiple-point statistics, which may be required in two distinct cases.

First, consistency problems may arise when the training image does not have exactly the same one-point and two-point distribution as the underlying phenomenon. This consistency problem occurs frequently, since training images usually are not available with the exact same histogram of the variable that is being modelled (proportions of facies). Direct use of the multiple-point statistics extracted from the training image will distort the histogram of the simulated realizations generating a bias in the proportions. A basic requirement for the simulated model to be accepted as a plausible representation of the true phenomenon is that it honours a given histogram. Scaling of multiple-point statistics is then required to allow the use of the training image in a consistent manner with the data and preserving its character, hence generating acceptable numerical models.

Another relevant use of a scaling procedure for multiple-point statistics is to impose non-stationary features, but preserving the character provided by the training image. One could devise the use of a single training image to model a field with locally varying proportions of the facies.

We present a methodology for scaling multiple-point statistics to generate consistent results from simulation methods that account for this information and to consider non-stationary features during the modeling process. The methodology is general and could be adapted to be used with continuous variables, as long as the data are coded as indicators by defining classes through a set of thresholds. In what follows, the problem is presented in the case of a categorical variable, which is where simulation accounting for multiple-point statistics has seen a faster development.

Problem setting

Consider a training image that depicts the spatial arrangement of K categories or facies. The global proportions with which these categories are present in the training image are denoted: $p_k, k = 1,...,K.$

The general appearance of the training image is deemed appropriate to model a given geological setting, hence the modeler decides the training image is to be used for inference of the multiple-point statistics that a given simulation algorithm will impose to a set of realizations.

The multiple-point statistics to be considered are defined by a spatial arrangement of nodes and by the combination of facies values in these nodes. If we consider the case, where an N-points

statistic is considered, then K^N possible combinations are available. Each of these combinations occurs in the training image with a given frequency. In fact, as soon as N or K become relatively large, there is big chance that many of the K^N combinations do not occur in the training image.

Each one of the possible combinations is identified with an index that completely defines the facies values within the N points, however the ordering to identify the points in the pattern must be defined prior to the calculation of the index of each multiple-point event (Figure 3). The index of each multiple-point configuration is calculated as:

$$j = 1 + \sum_{n=1}^{N} (i_n - 1) \cdot K^{n-1}$$

where i_n is the code of the nth node of the pattern that identifies its facies. Here, it is assumed that facies are denoted by consecutive integers starting with code 1.

The frequency of each multiple-point event in the training image is denoted: $f_j, j = 1,..., K^N$. Knowing the indexing and the frequencies with which each multiple-point configuration occurs, permits calculation of the facies proportions $p_k, k = 1,..., K.$. Denoting by $p_{k,j}; k = 1,..., K; j = 1,..., K^N$, the proportion of facies k in the multiple-point arrangement identified with the index j, the proportions of the facies in the training image can be retrieved from the multiple-point statistics as follows. Consider the multiple-point index of interest is j. We can calculate the value of the nth node, by taking:

$$i_N = int\left(\frac{j-1}{K^{N-1}}\right)$$

where *int* represents the integer part of the division. Subsequently, the indexes of the n–1 remaining nodes of the pattern can be calculated recursively using the residual of the division (fractional part):

$$i_n = int\left(\frac{frac\left(\frac{j-1}{K^n}\right)}{K^{n-1}}\right) \qquad n = 1,..., N-1$$

Knowing the facies values of the N nodes of each multiple-point index and their frequencies, the univariate proportions can be easily calculated.

Now, the simulated model must honor a set of statistics inferred from a set of samples. For instance, the available data may show that the global declustered proportions of the facies in the domain are: $p_k^{\text{target}}, k = 1,..., K$, with p_k not necessarily equal to p_k^{target} for some $k = 1,..., K$.

The goal of this paper is to propose a methodology to calculate the corrected frequencies of multiple-point events that will honor the target proportions, preserving the "character" of the training image, that is, keeping the features that make it distinct. These corrected frequencies are denoted: $f_j^*, j = 1,..., K^N$.

Scaling approaches for multiple-point statistics

To understand the concept of scaling multiple-point statistics, the following example illustrates possible outcomes from an increase in the proportion of a facies in a binary case and where the

scaled models preserve the character of the original training image. A field of 1000 by 1000 pixels is populated with 10 by 10 pixels squares, where the centre of the squares are randomly located in the field, that is they are generated through a Poisson process (

Figure 4). There are enough squares to cover 20% of the domain, leaving the remaining 80% as background. Scaling this training image so that the proportion of squares goes up to 40% may generate two equally valid outcomes: we can have more 10 by 10 pixels squares or have a smaller number of larger squares, say 20 by 20 pixels squares. These situations are of course extremes and since we do not know the exact multiple-point statistics of the variable and we borrow this information from a training image that does not exactly match the proportions, a scaling procedure that lays somewhere in between the two extremes presented above can be used as a modelling decision to handle the problem.

Two approaches could be devised to scale multiple-point statistics when the univariate distribution changes:

- Scaling based on how a variable randomly distributed changes its multiple-point statistics when the univariate proportions of the facies change.
- Scaling based on how a categorical variable obtained by multiple truncations of a correlated multiGaussian continuous variable changes as the thresholds that define the facies are modified.

Dimensionality is always a problem when dealing with multiple-point statistics since the combinatorial becomes rapidly very large as the number of points in the multiple-point configuration N or the number of facies K increase. For example, 10 facies can be arranged on a 9 points pattern in 1 billion possible combinations. Of course, this example is extreme, but pattern size often increases in an exponential fashion: 4, 9, 16, 25 points in 2D, or 8, 27, 64, 125 points in 3D. This problem may be partially solved by proceeding pairwise, that is, considering the most relevant facies and all other facies together. Then, freezing the most important facies already simulated, one can simulate another relevant facies against all remaining facies on a reduced domain, and so on, until all facies have been individually taken into account.

Scaling MPS based on the random case

The first approach proposed is based on how multiple-point statistics change when the facies codes of the points in the pattern are randomly distributed. In this case, the frequency with which a multiple-point event occurs can be calculated as the product of the probabilities of occurrence (proportions) of each facies values in its nodes. Since the facies values are considered uncorrelated, it is straightforward to scale the frequency of occurrence of multiple-point statistics. The frequency of a multiple-point configuration can be calculated by:

$$f_j = \prod_{k=1}^{K} p_k^{N \times p_{k,j}}$$

Since $p_{k,j}$ is the proportion of facies k in class j, $N \times p_{k,j}$ is the number of occurrences of facies k in class j. For example (Figure 5), considering a case where two facies are available and the probability of facies 1 prevailing at a given location is 0.25, then the probability of having a four-points configuration where facies 1 prevails in two nodes and facies 2 prevails in the remaining nodes would be calculated as:

$$f_{13} = 0.25^{4 \times 0.5} \cdot 0.75^{4 \times 0.5} = 0.03515625$$

Scaling the random case is quite simple; multiplying the original frequency by a series product of the desired frequency over the old frequency results in a simple equation:

$$f_j^* = f_j \times \prod_{k=1}^{K} \left(\frac{p_k^{\text{target}}}{p_k} \right)^{N \times p_{k,j}}$$

It is easy to see that the scaled multiple-point frequencies honour the target univariate proportions:

$$f_j^* = f_j \times \prod_{k=1}^{K} \left(\frac{p_k^{\text{target}}}{p_k} \right)^{N \times p_{k,j}} = f_j \times \frac{\prod_{k=1}^{K} \left(p_k^{\text{target}} \right)^{N \times p_{k,j}}}{\prod_{k=1}^{K} \left(p_k \right)^{N \times p_{k,j}}} = f_j \times \frac{f_j^*}{f_j} = f_j^*$$

Following up on the previous example, we can see that a change in the global proportions of the facies will change the probability of occurrence of the multiple-point event depicted in Figure 5. Since the uncorrelated case is simple, we can easily calculate this probability for the new global proportions. Consider the case, the probability of facies 1 prevailing at a location goes up to 0.4, leaving a probability of 0.6 for facies 2. Considering the same four-points configuration, we can calculate its new probability of occurrence:

$$f_{13}^{\text{new}} = 0.4^{4 \times 0.5} \cdot 0.6^{4 \times 0.5} = 0.0576$$

It can be checked that the expression provided for scaling multiple-point statistics in the random case provides the same result calculated above:

$$f_{13}^* = f_{13} \times \left(\frac{p_1^{\text{target}}}{p_1} \right)^{N \times p_{1,13}} \times \left(\frac{p_2^{\text{target}}}{p_2} \right)^{N \times p_{2,13}}$$

$$= 0.03515625 \times \left(\frac{0.4}{0.25} \right)^{4 \times 0.5} \times \left(\frac{0.6}{0.75} \right)^{4 \times 0.5} = 0.576$$

This expression is valid only in the case of a multiple-point statistics from an uncorrelated variable (pure nugget effect). As soon as the variable is spatially correlated, multiple-point proportions cannot be directly calculated. An iterative approach is proposed next to calculate the multiple-point frequencies that honour the target proportions, from the initial multiple-point frequencies inferred from the training image.

The formula above overcorrects the multiple-point frequencies. Convergence can be achieved by iterating using the following modified expression:

$$f_j^* = f_j \times \prod_{k=1}^{K} \left(\frac{p_k^{\text{target}}}{p_k} \right)^{p_{k,j}}$$

Multiple-point frequencies for all indexes $j = 1,...,K^N$ must be updated. The new global proportions $p_k, k = 1,...,K$ are recalculated for each iteration. This formula does not require a large number of iterations, and usually, the desired multiple-point frequencies can be obtained with a nearly perfect match of the target proportions, with less than 50 iterations, which takes only a few seconds.

Scaling MPS based on the multiGaussian case

An alternative method consists on considering how a categorical variable defined by multiple truncations of a multiGaussian field changes as the thresholds for truncations change. The idea can be understood quite simply, by considering a two-point pattern (

Figure 6). A generalization to N points can be achieved, associating a multivariate Gaussian distribution of dimension N.

In this simple example, two points separated by a vector **h** constitute the multiple-point pattern. Initially, we can consider a binary variable, hence it can be defined by a single truncation of the Gaussian variable with a threshold y_C. The two categories (facies) have probabilities:

$$p_1 = G(y_C)$$
$$p_2 = 1 - p_C = 1 - G(y_C)$$

An indicator variogram can be calculated and associated with a variogram for the underlying multiGaussian variable (Kyriakidis and others, 1999). Given the spatial arrangement of the two points, Their spatial correlation can be calculated from the variogram model and assuming multivariate Gaussianity, all conditional probabilities can be retrieved. The probabilities of all four possible multiple-point events can be calculated, by using conditional probabilities that only depend on two-point statistics, that is, the variogram model. For example, assuming facies 1 is associated with the continuous Gaussian variable when the value is below the threshold y_C:

$$\text{Prob}(i_1 = 1; i_2 = 2) = \int_{-\infty}^{y_C} \int_{y_C}^{\infty} g(y_1, y_2)\, dy_2\, dy_1$$

where $g(y_1,y_2)$ is the bivariate Gaussian probability density function.

This simple case can be easily extended to multiple points ($N > 2$), however, calculating the joint probabilities becomes cumbersome as the dimensionality increases. A simple numerical approximation can be used to calculate the multiple-point event frequencies.

In the multiple-point case, all correlations between the nodes in the pattern must be taken into account in order to approximate the integration of the multivariate Gaussian distribution using the appropriate thresholds to define the facies. A Monte-Carlo approach to simulate multivariate Gaussian vectors using the matrix decomposition method is proposed (Davis, 1987). The general procedure is:

1. Given the training image and target proportions $p_k, k = 1,...,K$ and $p_k^{\text{target}}, k = 1,...,K$, calculate the thresholds that give those probabilities when truncating a standard Gaussian distribution:

$$y_{C,1} = G^{-1}(p_1) \qquad y_{C,1}^{\text{target}} = G^{-1}(p_1^{\text{target}})$$
$$\vdots \qquad\qquad \vdots$$
$$y_{C,N-1} = G^{-1}\left(\sum_{i=1}^{N-1} p_i\right) \qquad y_{C,N-1}^{\text{target}} = G^{-1}\left(\sum_{i=1}^{N-1} p_i^{\text{target}}\right)$$

2. Compute the indicator variograms and infer the variogram of the Gaussian variable that better fits the calculated indicator variograms.
3. Given the separation vectors between the N points in the multiple-point pattern, build the N by N matrix of covariances between points in the pattern (

4. Figure 7):

$$C = \begin{bmatrix} C(\mathbf{h}_{11}) & C(\mathbf{h}_{12}) & \cdots & C(\mathbf{h}_{1N}) \\ C(\mathbf{h}_{21}) & C(\mathbf{h}_{22}) & \cdots & C(\mathbf{h}_{2N}) \\ \vdots & \vdots & \ddots & \vdots \\ C(\mathbf{h}_{N1}) & C(\mathbf{h}_{N2}) & \cdots & C(\mathbf{h}_{NN}) \end{bmatrix}$$

5. Using the matrix decomposition simulation algorithm generate a large number of simulated multivariate vectors that honor the covariance between the points. The algorithm proceeds as follows:
    a. Decompose the covariance matrix as a product of a lower (**L**) and an upper (**U**) triangular matrix such that $\mathbf{L} = \mathbf{U}^{\mathbf{T}}$ (Cholesky decomposition).
    b. Draw a vector of independent standard Gaussian deviates **w**.
    c. Generate a vector of correlated Gaussian values by multiplying the lower matrix by the vector of random normal deviates: $\mathbf{y} = \mathbf{L}\,\mathbf{w}$.
6. Identify the multiple-point index of each simulated vector by applying the truncation rule defined by the thresholds, that is transform the vectors of correlated Gaussian values to vectors of facies values, by coding the values as facies using the thresholds. Then, calculate the multiple-point index corresponding to each configuration of facies.
7. For each multiple-point configuration that existed in the training image, calculate the proportion of simulated vectors that match that multiple-point index when applying the thresholds calculated for the training image proportions and when the target proportions are used.
8. Correct the multiple-point statistics of the training image, using a ratio of those proportions.
9. Standardize to ensure the sum of proportions from the scaled multiple-point frequencies add to one.

This procedure accounts for the correlation between the points. It is an approximation due to the fact that the Gaussian variable is continuous, hence all multiple-point combination of indexes in the pattern has a probability of occurrence, which is not the case in the training image, since this is a discrete representation. A correction is required to match exactly the target proportions, but this cannot always be achieved, since using the multivariate Gaussian case imposes constraints with respect to the occurrence of "symmetric" events: for instance the presence of facies one on the right of facies two implies that facies two must exist also on the right of facies one. For this reason, this approach is generally not recommended.

Examples

First, to illustrate the use of these scaling approaches, the small example presented in

Figure 4 is followed up. The multiple-point histogram for a two by two points pattern is calculated. The frequency of occurrence of all 16 multiple-point events is computed and a histogram of the frequency with which each indexed event occurs is plotted as a summary. The 16 possible events can be easily interpreted in terms of the geometry of the objects (squares), as shown in Figure 8.

The reference image (map on the left hand side of

Figure 4) is used to calculate the multiple-point frequencies in a 2 by 2 pixels pattern and these are scaled to reach a global proportion of squares of 0.40. These scaled statistics are then compared with the two extreme cases presented on the right hand side of

Figure 4. Figure 9 shows the three multiple-point histograms superimposed. It can be seen that the scaling using the random approach provides statistics that lie in between the two extreme cases. We can expect that these scaled up statistics represent a case where there are more squares in the domain, but these are also slightly larger than the ones in the training image. The "relevant" features of the reference image are then captured in the statistics.

A second example has been designed to illustrate the use of the scaling procedure proposed. A binary two dimensional training image has been created with black horizontal stripes two units thick in a white background. The training image is shown in Figure 10. The original proportions are $p_{black} = 0.25$ and $p_{white} = 0.75$.

Multiple-point statistics have been inferred from the training image and unconditional realizations have been computed for different proportions of black (stripes) and white (background) facies. A simulated annealing program to match the multiple-point statistics was prepared. Different pattern sizes imply control over different scales of the phenomenon. For example, Figure 11 shows two realizations considering a 4 by 4 nodes pattern. Notice that with this pattern size the general appearance of the resulting models is close to the training image. If a different pattern is considered, the long range features may not be properly captured. Figure 12 shows realizations made with different pattern sizes to illustrate this effect.

The scaling procedure was used to scale up the proportion of the black facies. The initial proportions (25% black, 75% white) was changed and the multiple-point statistics were scaled. These were then used to simulated realizations with the intent of generating realizations that kept the essential features of the training image, but honouring the facies proportions imposed to the new models.

Figure 13 shows three realizations constructed using the multiple-point statistics for a 3 by 3 nodes pattern, scaled to match the following facies proportions: 30% black and 70% white. These proportions were then changed to 50% for the black facies and 50% for the white background, and realizations with statistics from 2 by 2, 3 by 3, and 4 by 4 nodes patterns considered. These results are shown in Figure 14.

It can be seen from the realizations displayed above that, depending on the pattern size, two situations can arise: the stripes may become wider if the pattern is small enough not to capture its entire width, or the stripes may become more abundant if the pattern is big enough to capture the relationships between the width of the black facies objects and the background, as is the case when 3 by 3 or 4 by 4 patterns are used.

Conclusions

Scaling multiple-point statistics can improve the current use of multiple-point geostatistical simulation techniques, by allowing locally varying proportions of facies to be reproduced still honouring the relationships captured by multiple-point statistics. Another important use of a multiple-point scaling technique is the use of a representative training image that does not exactly match the target simulation proportions.

This article presents an approach to scale multiple-point statistics by iteratively changing them in a manner similar to the change that occurs in multiple-point statistics of a purely random field where the proportions of the facies are modified. This approach can be shown to provide

reasonable values of the statistics that lie in between two extreme cases that are easy to show in the case of objects: an increase in the number of objects or an increase in the size of the objects.

The scaling procedure has the potential of being incorporated in a non-stationary simulation algorithm that honors the features provided by a training image, through the local updating of multiple-point statistics.

Acknowledgements

The authors would like to acknowledge the industry sponsors of the Centre for Computational Geostatistics at the University of Alberta and the Chair in Ore Reserve Evaluation at the University of Chile sponsored by Codelco Chile for supporting this research.

References

- Caers, J. and Journel, A. G., Stochastic reservoir simulation using neural networks trained on outcrop data, in 1998 SPE Annual Technical Conference and Exhibition, New Orleans, LA. Society of Petroleum Engineers. SPE paper # 49026, 1998, p. 321-336.
- Davis, M. W., Production of conditional simulations via the LU triangular decomposition of the covariance matrix. Mathematical Geology, vol. 19, no. 2, 1987, p. 99-107.
- Deutsch, C. V., Annealing Techniques Applied to Reservoir Modeling and the Integration of Geological and Engineering (Well Test) Data. PhD thesis, Stanford University, Stanford, CA, 1992.
- Guardiano, F. and Srivastava, M., Multivariate geostatistics: Beyond bivariate moments, in A. Soares, editor, Geostatistics Tróia'92. Kluwer, Dordrecht, vol. 1, 1993, p. 133-144.
- Kyriakidis, P. C., Deutsch C. V., and Grant M. L., Calculation of the normal scores variogram used for truncated Gaussian lithofacies simulation: theory and FORTRAN code. Computers & Geosciences 25 (2): 161-169, March 1999.
- Ortiz, J. M., Characterization of High Order Correlation for Enhanced Indicator Simulation. PhD thesis, University of Alberta, Edmonton, AB, Canada, 2003.
- Ortiz, J. M. and Deutsch, C. V., Indicator simulation accounting for multiple-point statistics. Mathematical Geology, vol. 36, no. 6, 2004, p. 545-565.
- Ortiz, J. M. and Emery, X., Integrating multiple-point statistics into sequential simulation algorithms, in O. Leuangthong and C. V. Deutsch, editors, Geostatistics Banff 2004. Kluwer Academic, Dordrecht, The Netherlands, Vol. 2, 2005, p. 969-978.
- Strebelle, S. and Journel, A. G., Sequential simulation drawing structures from training images, in 6th International Geostatistics Congress, Cape Town, South Africa, Geostatistical Association of Southern Africa, 2000.

Figure 1: an exhaustive image showing intrincated relationships of four facies in multiple-point patterns that cannot be easily captured by conventional simulation techniques.

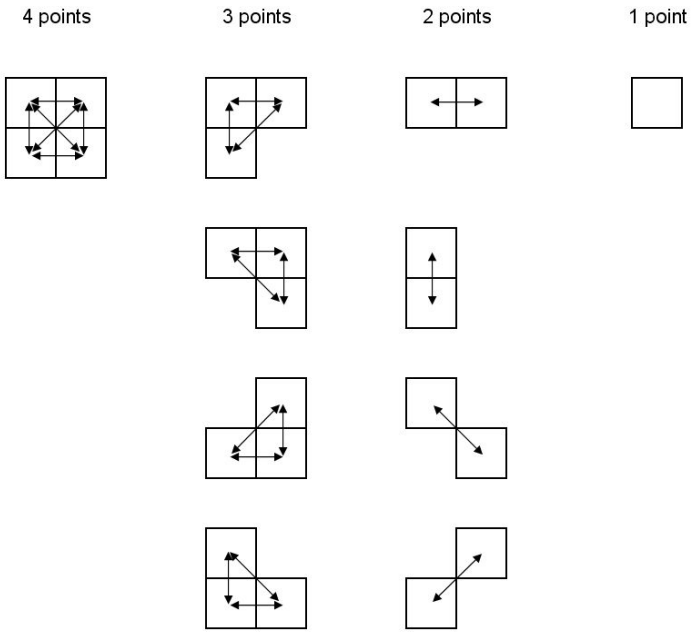

Figure 2: A four-points pattern and all lower order configurations that are implicitly matched by honouring the four-point statistics: three-points statistics in four configurations; two-points statistics for the corresponding lag distances separating the centres of the nodes (these represent indicator variogram values); and the one-point statistics that corresponds to the histogram. Nodes are represented by their surrounding squares for illustration.

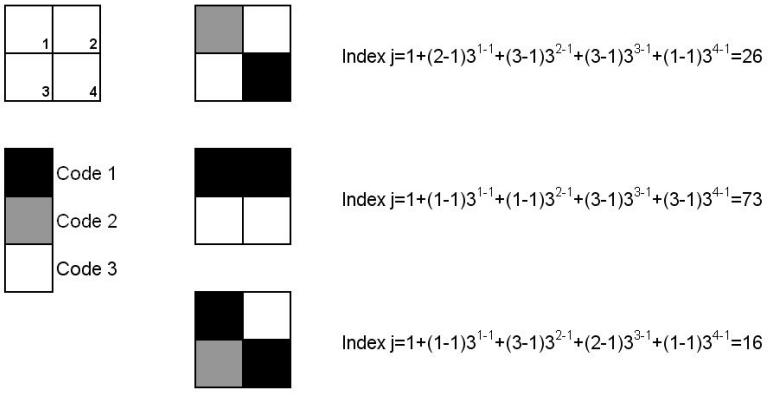

Figure 3: A four-point configuration, the order for considering the nodes and codes of the facies. Three examples of calculation of the multiple-points indexes are illustrated.

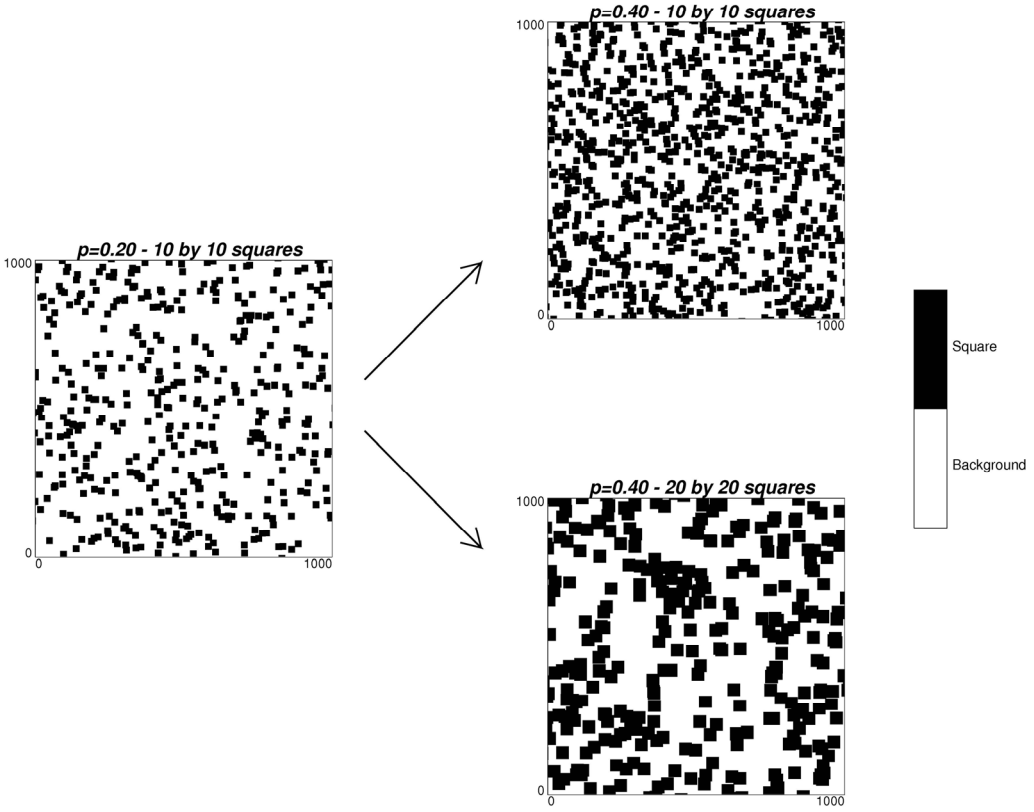

Figure 4: Two possible upscaled models from the training image on the left. The top right map shows more squares of the same dimension as in the reference image; the bottom right map shows larger squares. Both maps have a proportion of pixels belonging to squares of 0.40.

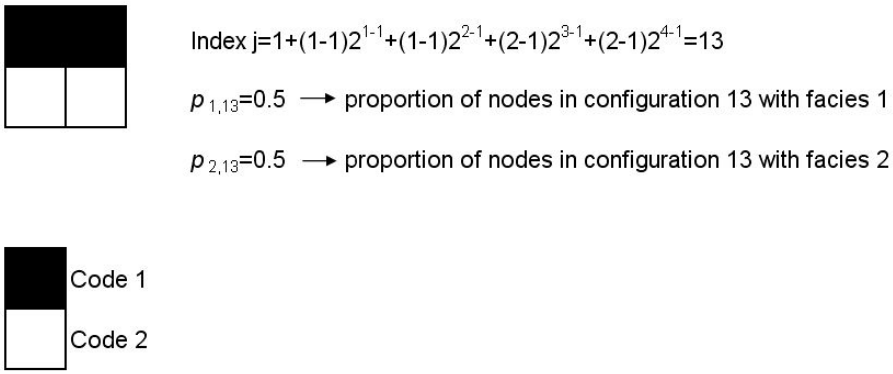

Figure 5: example of four-points configuration with two facies to calculate the probability of occurrence of an event in the random case.

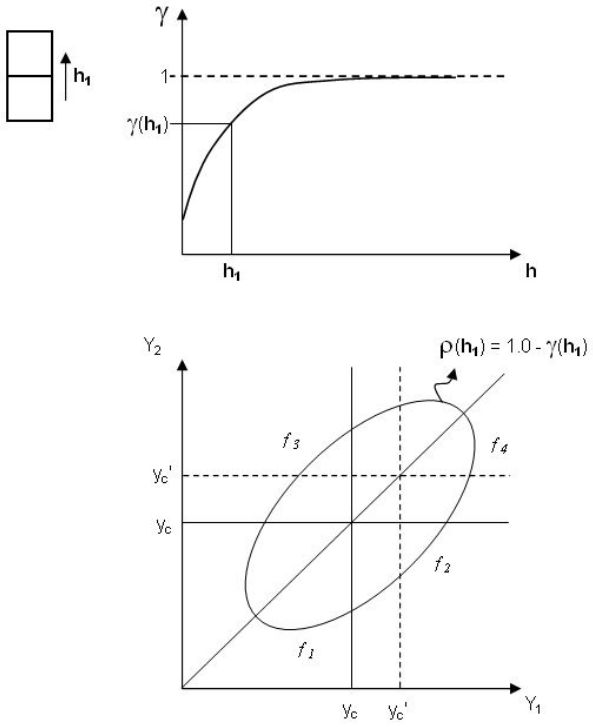

Figure 6: An illustration of the relationship between global proportions, defined by the thresholds that truncate the multiGaussian distribution, and the multiple-point frequencies. A two-points pattern of nodes separated by a lag vector h allows calculation of the correlation coefficient from the normal scores variogram γ(h). This defines the bivariate Gaussian distribution that is truncated so that the probability of being below y_C is p_1 and completely defines the multiple-point probabilities. These probabilities change in a tractable manner when the global proportions are changed by modifying the threshold.

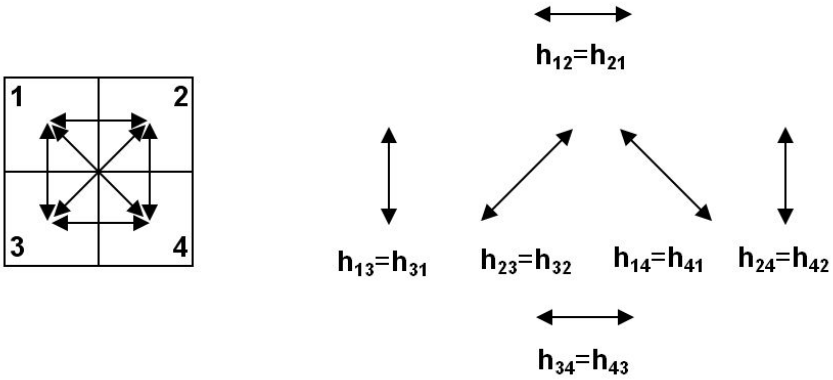

Figure 7: Four-points configuration and the corresponding lag separation vectors used to build the covariance matrix to simulate multivariate Gaussian vectors.

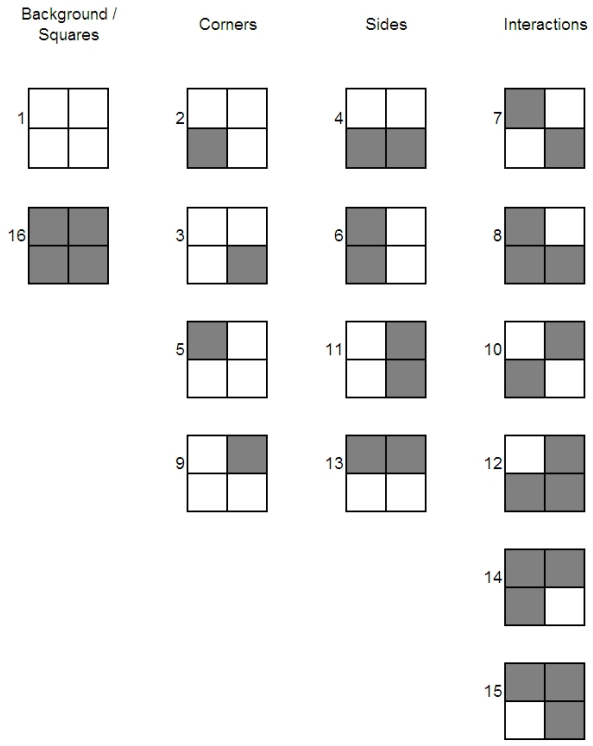

Figure 8: Geometric interpretation of the different multiple-point configurations for the example of drawing squares randomly in the domain.

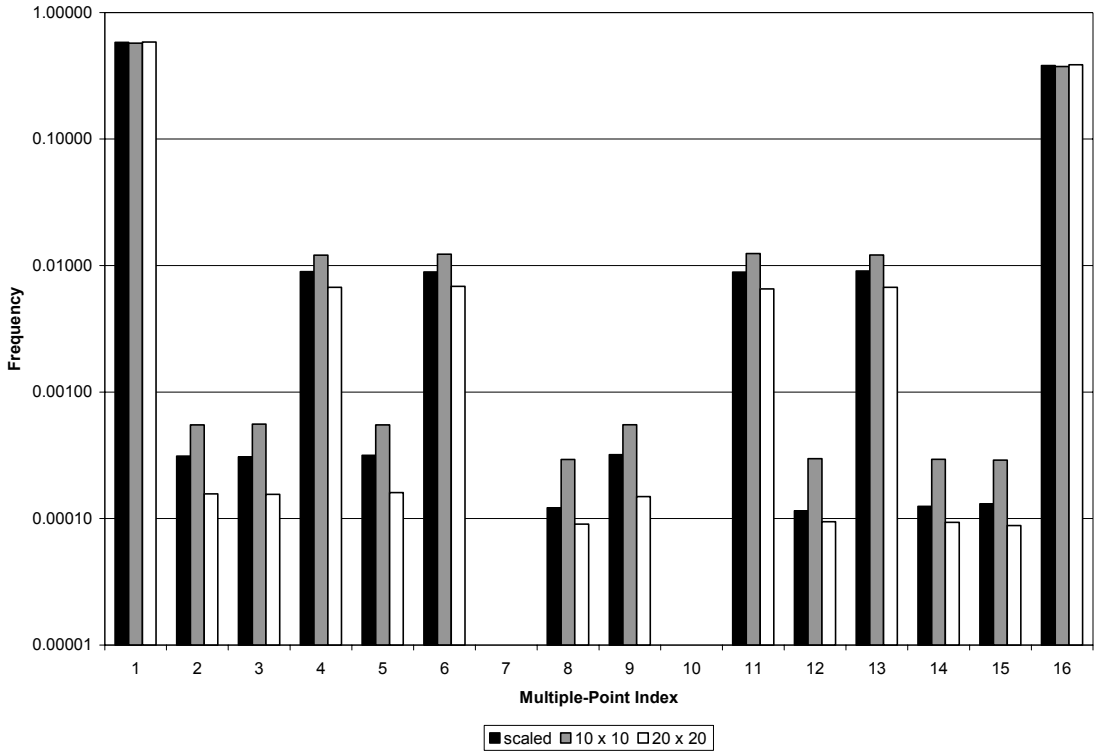

Figure 9: Multiple-point histogram for the scaled statistics and the two extreme cases: increased number of 10 by 10 squares and increased size of the squares to 20 by 20. A logarithmic scale has been used for the frequency axis to better show that he scaled statistics lie in between the frequencies for the extreme cases.

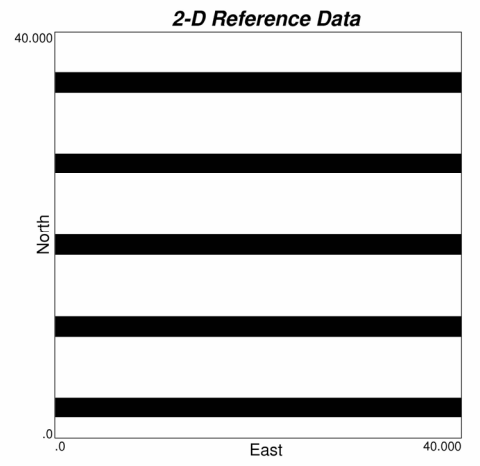

Figure 10: Reference training image. Black stripes are two nodes wide and separated by 6 nodes in the north direction.

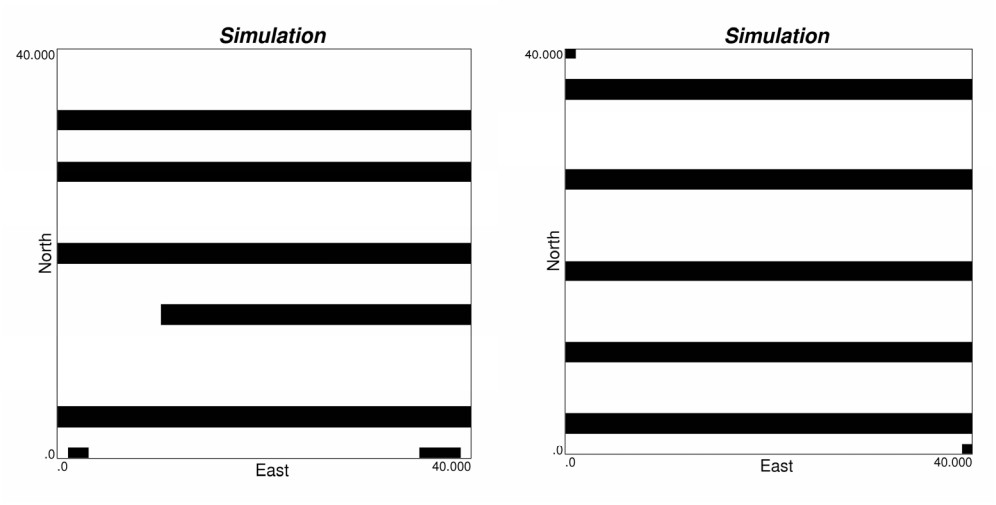

Figure 11: Two realizations that honor the original proportions, made considering a 4 by 4 nodes pattern.

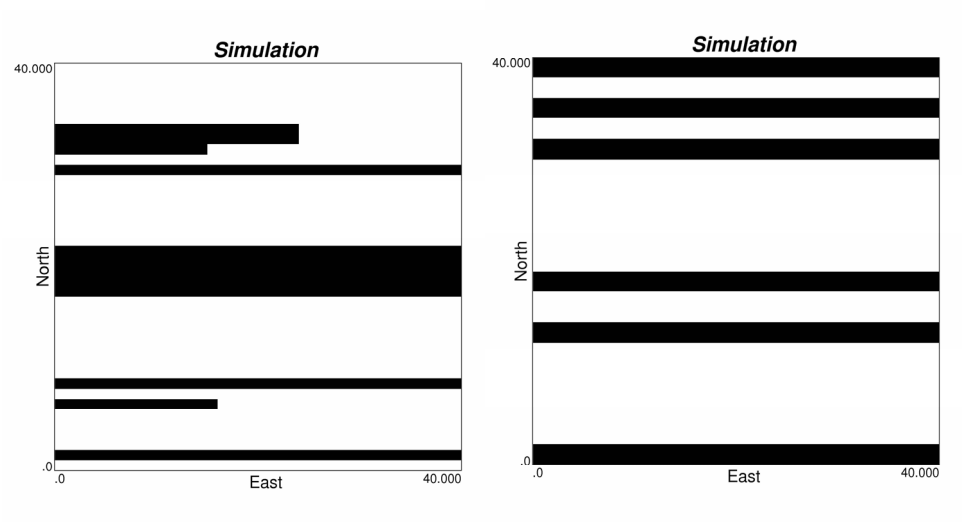

Figure 12: Two realizations that honor the original proportions, made considering statistics extracted from a 2 by 2 (left) and a 3 by 3 nodes pattern (right).

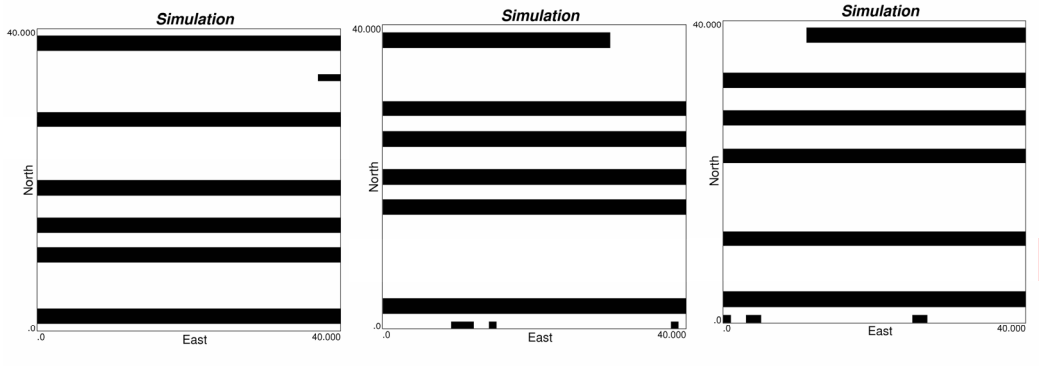

Figure 13: Three realizations considering the statistics from a 3 by 3 nodes pattern and scaled proportion of 30% for the black facies and 70% for the white background.

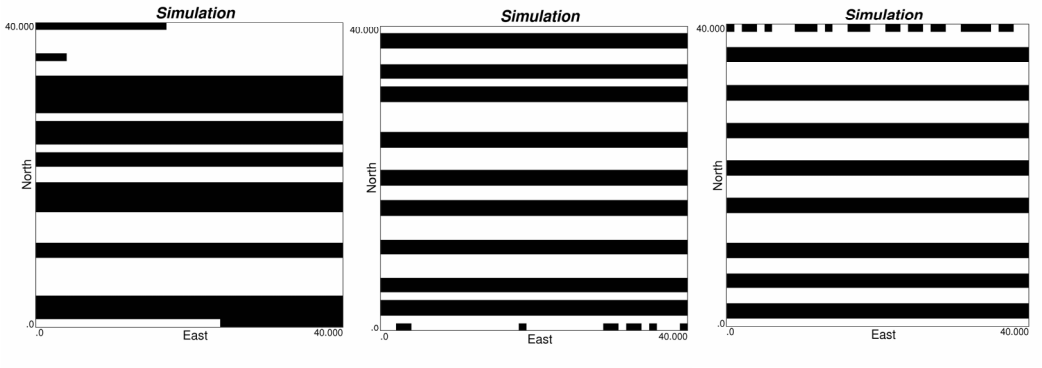

Figure 14: Three realizations with multiple-point statistics scaled to 50% black facies and 50% white background, with increasing pattern size: 2 by 2 (left), 3 by 3 (middle), and 4 by 4 (right).